KARYA ILHAM PUBLISHING
CONNECTING MINDS, BRIDGING IDEAS, INSPIRING READERS

Journal of Advanced Research in Computing and Applications

Journal homepage:
https://akademiabaru.com/submit/index.php/arca/index
2462-1927

Hybrid Deep Learning for Sentiment Analysis: A Bibliometric Perspective on Research trends And Gaps

Shakirah Mohd Sofi[1,2,*], Ali Selamat[2,3], Zatul Alwani Shaffiei[2]

[1] Department of Computing, Faculty of Creative Multimedia and Computing, Selangor Islamic University, Bandar Seri Putra, Kajang 43000, Selangor, Malaysia
[2] Malaysia-Japan International Institute of Technology (MJIIT), University Technology Malaysia Kuala Lumpur, Jalan Sultan Yahya Petra, Kuala Lumpur 54100, Malaysia
[3] Faculty of Computing, Faculty of Engineering, & Media and Games Center of Excellence (MagicX), University Technology Malaysia, Skudai 81310, Johor Bahru, Malaysia

ARTICLE INFO

Article history:
Received 25 October 2025
Received in revised form 29 November 2025
Accepted 10 December 2025
Available online 17 December 2025

Keywords:
Hybrid Deep Learning; Sentiment Analysis; Bibliometric Analysis; Transformer Models; Research Trends; Cross-Domain

ABSTRACT

This bibliometric study examines hybrid deep learning architectures for sentiment analysis by analysing 18,373 publications (2010-2025) from Scopus through descriptive analysis, citation metrics, and keyword co-occurrence methods. Hybrid architectures mix different deep learning components, with ensemble methods and transformer-based models emerging as the two primary approaches researchers use. Ensemble methods showed up most frequently at 35%, but transformers grew fastest, increasing from 2.4% to 20.7% during the study period. China published the most with 5,909 papers, India came in second with 3,949, and the United States had 1,523. Across all the research, 12 main themes stood out. Ensemble methods appeared 3,675 times and transformers appeared 3,200 times across the papers reviewed. Among specific techniques, BERT variants grew from nearly invisible (0.1%) to 10.4%, while attention mechanisms climbed from 1.9% to 14.0%. Some areas got much less attention, though. Low-resource languages grew only 2.1%, domain adaptation just 0.3%, and real-time processing 2.4%, all falling far behind mainstream methods. These findings document how hybrid deep learning in sentiment analysis has evolved while identifying underexplored areas that present opportunities for methodological innovation and broader practical application.

1. Introduction

Social media, e-commerce sites, and online forums now generate massive amounts of user content, offering fresh ways to understand what people think about products, services, and events. Sentiment analysis, often called opinion mining, has become a crucial NLP tool that helps computers automatically identify and measure opinions and emotions in text. Businesses, policymakers, and researchers now depend on it to extract actionable insights from massive amounts of unstructured text.

* *Corresponding author.*
E-mail address: syakirah@uis.edu.my

https://doi.org/10.37934/arca.41.1.244260

Early sentiment analysis efforts depended primarily on traditional machine learning approaches such as Logistic Regression, Support Vector Machines, Naïve Bayes, Random Forest, and Decision Trees. These techniques worked reasonably well on structured data, especially when researchers paired them with feature extraction approaches such as Bag-of-Words, TF-IDF, and word embeddings like Word2Vec and GloVe. Deep learning changed everything, though. Architectures like Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM) networks, Bidirectional LSTM (BiLSTM), Gated Recurrent Units (GRUs), and transformer models such as BERT, RoBERTa, and GPT proved far better at automatically pulling out meaningful features and understanding complex semantic relationships [1]. Lately, hybrid deep learning frameworks have taken centre stage by combining multiple algorithms and architectures. Examples include CNN-LSTM, CNN-BiLSTM with attention mechanisms, BERT-CNN, BERT-LSTM, hierarchical attention networks, and ensemble strategies that bring together different approaches to tap into their complementary strengths [2]. These hybrid models have delivered strong results across various sentiment analysis tasks by merging CNNs' feature extraction abilities, RNNs' sequential processing power, and transformers' contextual grasp.

Despite the growing body of research on deep learning architectures for sentiment analysis, no comprehensive bibliometric analysis has yet systematically mapped how this field has evolved, what trends have emerged, and who the key contributors are. Researchers have proposed many different architectures, but we lack clarity on which ones have gained real traction, how themes have changed over time, what patterns exist across countries and institutions, and where opportunities lie. This makes it hard for researchers and practitioners to spot promising directions or grasp the bigger picture of what's happening in sentiment analysis.

In this paper, a comprehensive bibliometric analysis is conducted of the hybrid deep learning model related sentiment analysis research from 2010 to 2025.The research questions (RQs) that are addressed are as follows:

RQ1: What are the current publication and citation trends in hybrid deep learning for sentiment analysis from 2010 to 2025?

RQ2: Which hybrid deep learning architectures are most frequently employed in sentiment analysis research?

RQ3: Who are the most influential authors, institutions, and countries contributing to this field?

RQ4: What are the emerging themes and methodological trends in hybrid deep learning for sentiment analysis?

RQ5: What novel techniques and architectural innovations have gained traction in recent years?

The remainder of this paper is organized as follows. In Section 2, the theoretical foundations of sentiment analysis and hybrid deep learning models are reviewed. Section 3 describes the procedures through which the data were collected and the bibliometric analysis was carried out. Section 4 presented the results and discussion including publication trends, citation patterns, and thematic analysis. The concluding section, Section 5, brings together findings from each research question and discusses what they reveal for theory and practice.

This bibliometric analysis helps several different audiences understand sentiment analysis research better. Researchers can see how the field grew and changed, which studies mattered most, what directions became popular, and where work still needs doing. Industry practitioners find out which methods and architectures delivered results in practice, guiding their technology choices. Academic institutions and funding bodies discover patterns in where research takes place and how scientists work together across borders, useful knowledge for planning funding and collaborations.

The analysis maps what sentiment analysis research has accomplished and where it falls short, giving the research community clearer direction for future work.

2. Literature Review

2.1 Theoretical Foundations of Sentiment Analysis

Sentiment analysis, often called opinion mining, examines the opinions, feelings, evaluations, attitudes, and emotions people express about products, services, organizations, individuals, issues, events, and their features [3]. Social media has exploded with information in recent years, making sentiment analysis crucial for understanding commercial outcomes, political developments, public security concerns, and broader social perspectives [4].

Traditional approaches relied on lexicon-based methods and classical machine learning algorithms like Support Vector Machines, Naive Bayes, and Decision Trees [5]. While reasonable performance was achieved by these methods on structured datasets, extensive manual feature engineering was required, and significant difficulties were encountered when handling linguistic complexities such as sarcasm, negation, and context-dependent meanings [6]. Furthermore, the "one-model-fits-all" assumption that underlies traditional approaches has been shown in recent research to be fundamentally restrictive when heterogeneous data sources are involved, as different data distributions necessitate tailored modelling strategies [7]. Traditional machine learning methods using LDA and TF-IDF for feature extraction work reasonably well for aspect-based sentiment analysis [8], but hybrid deep learning models do a better job of capturing the subtle, aspect-specific sentiments across different domains.

2.2 Deep Learning Approaches

Deep learning changed sentiment analysis by allowing models to learn features from raw text directly, eliminating the need for manual feature design [1]. While CNNs handle local patterns and n-gram features effectively, they face challenges with word order and grasping relationships over longer text spans [9], [10]. Sequential language is processed more effectively by LSTMs, through which context across sentences is maintained and long-range dependencies are captured, although this comes at high computational cost [11]. The sequential processing requirement prevents parallelization, creating a bottleneck whereby scalability is significantly limited for longer sequences [12]. Relevant text has been identified by attention mechanisms, leading to better performance [13]. Context understanding has been improved by BERT using self-attention [14]. Better emotion recognition has been achieved by RoBERTa [15].

2.3 Hybrid Deep Learning Models

Hybrid models integrate various methods and techniques, ensemble approaches included, to achieve better performance. Ensemble methods work by using several models together, where each model makes a prediction and then these predictions are combined by averaging or voting. Hybrid models are commonly adopted because different architectural strengths can be integrated, whereby the weaknesses of individual methods are effectively mitigated [2]. For aspect-based sentiment analysis, strong contextual representations are provided by BERT and subsequently refined through CNN layers, resulting in enhanced pattern recognition in BERT–CNN models [16]. In BERT–LSTM architectures, BERT provides rich contextual embeddings, while LSTMs are employed to model temporal dependencies, thus allowing longer and more complex texts to be processed effectively

[17]. Using TF-IDF feature weighting and a CNN+BiGRU architecture, achieved superior classification accuracy compared to conventional CNN, LSTM, and C-LSTM models in aspect-based sentiment analysis tasks [18].

The FakeStack, which combines BERT, CNN, and LSTM architecture to better understand context in text. Skip-convolution blocks are incorporated to retain essential information across the network, while deep CNN layers and LSTM units are jointly used to extract salient features and capture the progression of sentiment over time [19]. RoBERTa has been shown to outperform alternatives such as BERT, XLNet, and DistilBERT in the analysis of emotions in text. Remarkably, this study revealed that ensemble approaches use RoBERTa was combined with complementary models along with the integration of common-sense knowledge bases, significantly improved the system's capacity to generalize effectively across diverse textual contexts [20]. The increasing numbers of research devoted to hybrid deep learning in sentiment analysis reflects both promise of these methods and the evolving challenges in NLP.

3. Methodology

3.1 Data Collection and Data Processing

In this study, the Scopus database was selected for literature collection because of its extensive coverage of peer-reviewed journals, conference proceedings, and scholarly works across multiple disciplines. Search terms related to hybrid models, deep learning approaches and sentiment analysis contexts were developed and applied systematically. Keyword searches focusing on titles, abstracts, and keyword fields were used to search through titles, abstracts, and keyword fields, as shown in Figure 1. Inclusion criteria were defined and consistently applied during the selection process to maintain the quality and relevance of the collected literature. The search was limited to English-language publications, with document types selected for their scholarly value as described in Table 1. The distribution of source categories encompassing conference proceedings, journals, and books is presented in Table 2. Publications from 2010 to 2025 were examined to capture both foundational studies and recent advances in hybrid deep learning architectures for sentiment analysis. Through this systematic approach, 18,373 documents were retrieved on November 25, 2025, forming a comprehensive dataset that demonstrates the significant expansion and evolution of research in deep learning-based sentiment analysis over the fifteen-year period.

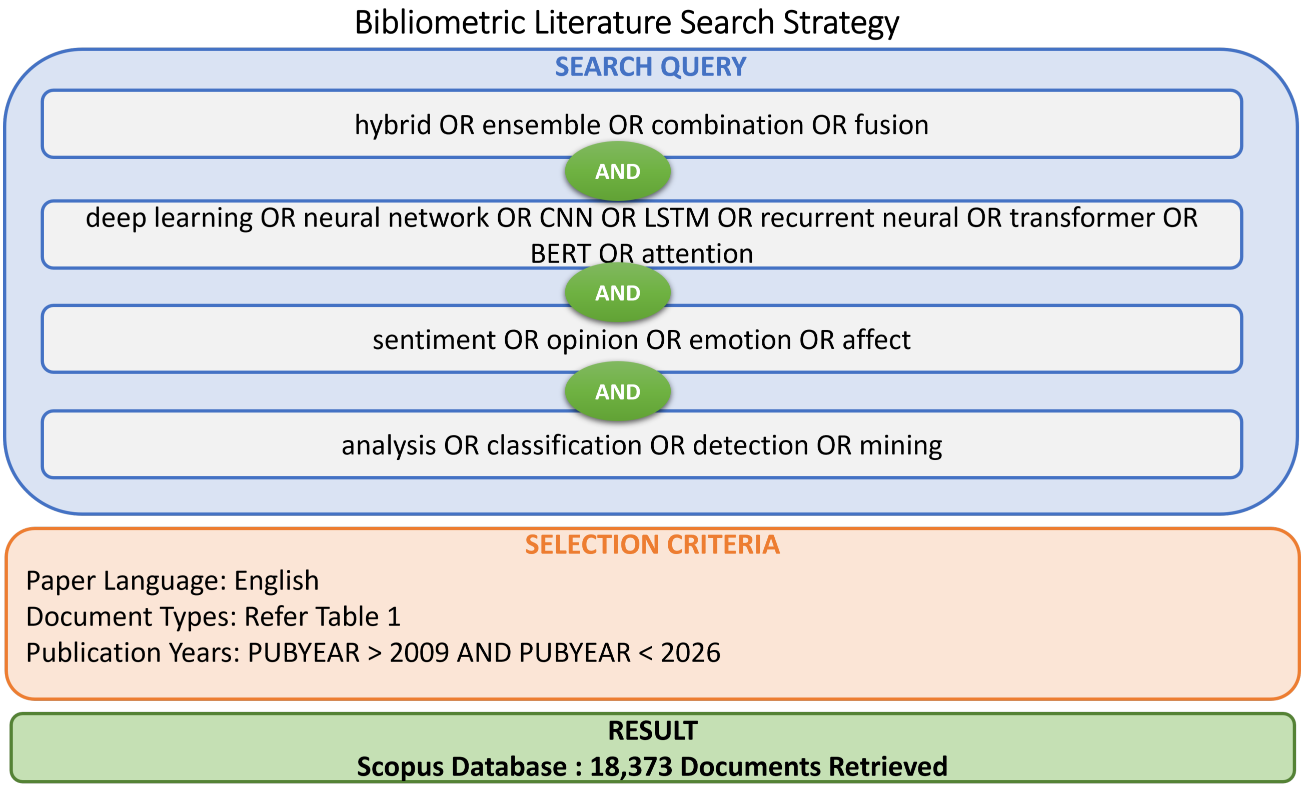

Fig. 1. Scopus data collection

*Retrieved Date: November 25, 2025

Table 1 Main Document Type

Document Type	Total Publication	Percentages (%)
Article	10071	54.81
Conference Paper	6168	33.57
Conference Review	1324	7.21
Review	454	2.47
Book Chapter	249	1.36
Retracted	45	0.24
Erratum	43	0.23
Book	19	0.10

Table 2 Main Source Type

Source Type	Total Publication	Percentages (%)
Journal	10683	58.15
Conference Proceeding	5418	29.49
Book Series	2107	11.47
Book	159	0.87
Trade Journal	6	0.03

3.2 Bibliometric Analysis Techniques

This study employed multiple bibliometric techniques to analyse research trends in hybrid deep learning approaches such as data analysis tools, bibliometric indicators and analytical techniques.

3.2.1 Data Analysis Tools

This bibliometric study was used analytical tools such as Harzing's Publish or Perish (PoP) and Python 3.9. Harzing's Publish or Perish extracted key citation metrics, while Python (with pandas, matplotlib and seaborn) supported data analysis, visualization, and network exploration.

3.2.2 Bibliometric Indicators

This study utilized Harzing's Publish or Perish and Python programming to examine citation patterns and research trends. Harzing's Publish or Perish calculated key bibliometric metrics like total citations, h-index, and g-index, showing which scholarly works had the most influence and how research areas grew over time. Python libraries then processed and visualized this citation data, Pandas handled data manipulation, NumPy performed calculations, and Matplotlib and Seaborn created visual representations. The synergy of these tools enabled a complete assessment of trends in hybrid deep learning applications for sentiment analysis.

3.2.3 Analytical Techniques

Five analytical techniques were methodically applied to explore the bibliometric data.

i. Descriptive analysis examined temporal trends in publication output from 2010 to 2025, geographic distribution patterns, publication sources, leading contributors by country and institution.
ii. Citation analysis was performed on the top most-cited articles to determine which works have been most influential and to understand citation patterns in the field.
iii. Keyword analysis identified the most frequent keywords and tracked their evolution over time to detect emerging research themes.
iv. Co-occurrence analysis on keywords appearing at the conceptual structure of the field and identifies thematic clusters, revealing which hybrid architectures and methodologies are most prominently researched.
v. Trend analysis was carried out using moving average calculations on publication data to identify growth patterns and to help project future research directions in hybrid deep learning.

4. Results and Discussion

4.1 Current Publication and Citation Trends

To address RQ1, publication and citation trends in hybrid deep learning for sentiment analysis were tracked from 2010 to 2025. Figure 2 displays the growth path of published research in hybrid deep learning for sentiment analysis. Various document types were considered and 18,373 papers were extracted from the Scopus database. A consistent upward trend in publication volume was observed, with substantial growth recorded between 2010 and 2025. Publications increased from 720 in 2021 to 2,273 in 2022, which represents a 12.37% gain.

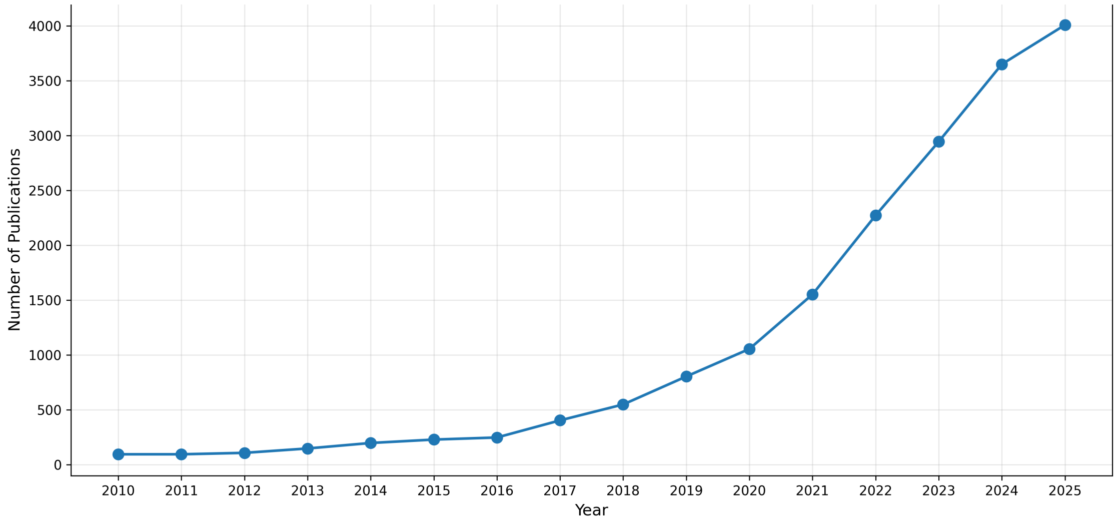

Fig. 2. The growth of published papers (Until 25 November 2025)

Table 3
Citation Metrics

Metrics	Data
Reference date	25 November 2025
Publication years	2010-2025
Citation years	15 (2010-20250
Papers	18373
Citations	266850
Citation / year	17790.00
Citations / paper	14.52
Citations / author	76954.00
Papers / author	5517.50
Authors / paper	3.75
Hirsch h-index	178
Egghe g-index	263

The citation metrics for papers retrieved up to November 25, 2025, are displayed in Table 3. Harzing's Publish or Perish software was employed to calculate these metrics from the Scopus database. Key indicators include total citations, yearly citation trends, average citations per paper, and citations per author. From the 13,873 publications examined in hybrid deep learning for sentiment analysis, 266,850 total citations were recorded, averaging 17,790 citations annually. Each paper received approximately 14.52 citations on average, and the overall collection achieved an h-index of 178 and a g-index of 263.

The exponential growth in publications and citation impact (h-index 178) indicates that hybrid deep learning approaches address critical needs in sentiment analysis and achieve higher accuracy and robustness than single model architectures. The 12.37% year-over-year increase shows that researchers continue to encounter complex sentiment analysis challenges like sarcasm detection, context-dependent meanings, and multi-lingual analysis that require innovative hybrid models. This sustained research momentum indicates that existing methods have not yet reached performance ceilings, which highlights the continued relevance of exploring novel architectural combinations.

4.2 The Hybrid Deep Learning Architectures

Table 4 and Figure 3 reveal that Hybrid-general or Ensemble methods (35%) and CNN-LSTM combinations (21%) dominate this field because researchers are more concerned with what really works than what is up to date. The reason is simple that sentiment analysis models encounter very different datasets and domains, so they need to be flexible enough to handle variety without starting over every time. Interestingly, transformer-based hybrids remain less popular (BERT-LSTM at 8%, RoBERTa at 6%) even though they perform well in theory. This reveals a real gap between what works in research papers and what practitioners can actually use. Transformers demand substantial computing resources and technical expertise, which explains their limited adoption in hybrid models. What researchers really need are practical approaches to leverage transformer strengths without the prohibitive computational expenses.

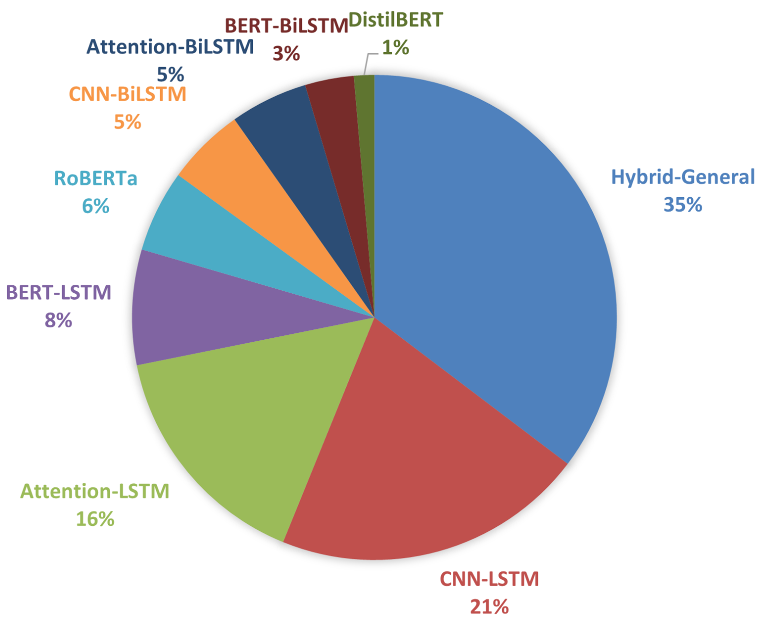

Table 4 Frequency of Architecture Appearing in Articles

Architecture	Frequent Appearances in Articles
Hybrid-General	1842
CNN-LSTM	1087
Attention-LSTM	818
BERT-LSTM	401
RoBERTa	285
CNN-BiLSTM	272
Attention-BiLSTM	270
BERT-BiLSTM	170
DistilBERT	71
XLNet	63
GRU-CNN	61
ALBERT	61
Transformer-CNN	42
BERT-CNN	33
BERT-GRU	2

4.3 Key Contributors: Authors, Institutions, and Countries

According to Table 5, Chinese authors have established a dominant position in the research field, with Wang, Y. standing out as the most productive scholar with 302 publications and 7,550 citations, succeeded by Zhang, Y. and Li, Y., whereby a consistent h-index of 25 has been achieved by all leading authors. This individual dominance is reflected at the national level in Figure 4, where China is shown to lead with 5,909 publications, which significantly surpasses India (3,949) and the United States (1,523). Figure 5 shows that Chitkara University in Punjab, India, is the most productive institution with 125 publications. It is followed by the Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology (64 publications) and the University of Chinese Academy of Sciences (60 publications). Tsinghua University and Beihang University, two well-known Chinese institutions, each made moderate contributions of 53 and 48 publications, respectively. This distribution pattern suggests that China's quantitative leadership in research output is evident at both individual and national levels, though the institutional diversity illustrated across the figures indicates a globally distributed research ecosystem where significant contributions are made by Indian and Chinese universities, thereby reflecting regional investments in this research domain and the collaborative nature of contemporary scientific inquiry.

Table 5
Top 10 Most Productive Authors

Rank	Author Name	Publications	Total Citations	h-index	Avg. Citations/Paper
1	Wang, Y.	302	7550	25	25
2	Zhang, Y.	276	6900	25	25
3	Li, Y.	257	6425	25	25
4	Li, X.	230	5750	25	25
5	Li, J.	226	5650	25	25
6	Liu, Y.	222	5550	25	25
7	Zhang, X.	214	5350	25	25
8	Wang, X.	204	5100	25	25
9	Wang, J.	196	4900	25	25
10	Wang, Z.	196	4900	25	25

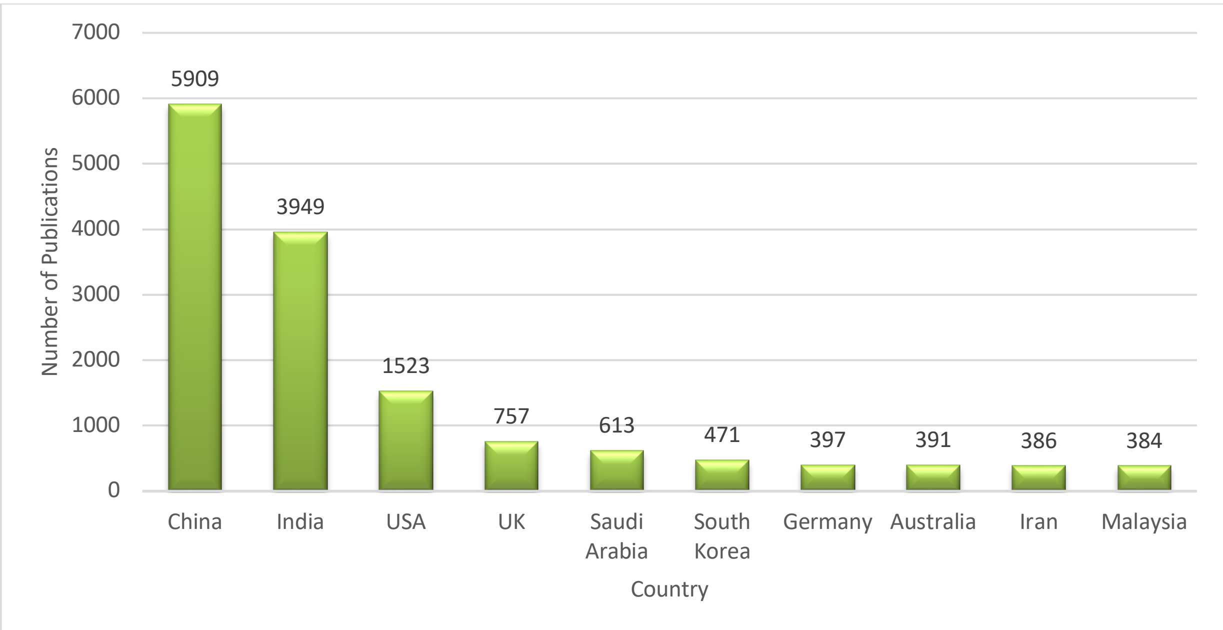

Fig. 4. Top 10 Contributing Countries

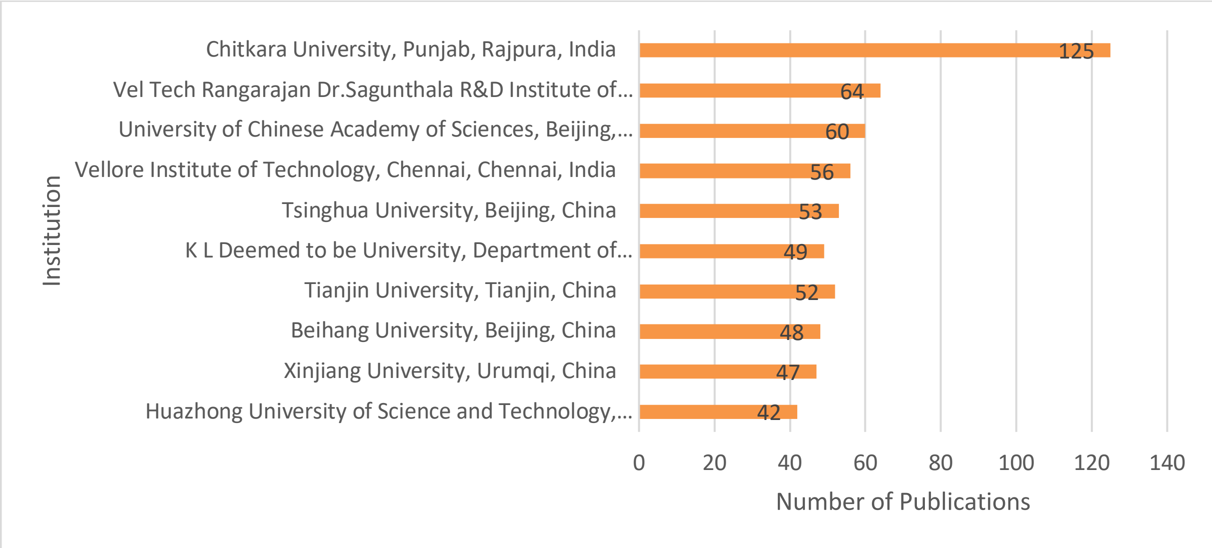

Fig. 5. Top 10 Contributing Institutions

4.4 Emerging Themes and Methodological Trends

Figure 6 shows common keywords in sentiment analysis research. "Sentiment analysis" and "deep learning" are the most frequent terms. "Emotion recognition" and "convolutional neural networks" appear regularly, reflecting their widespread adoption in the field. The presence of keywords like "multi-modal" and "social media" shows researchers have expanded beyond text to incorporate images and audio from social platforms in their emotion detection work.

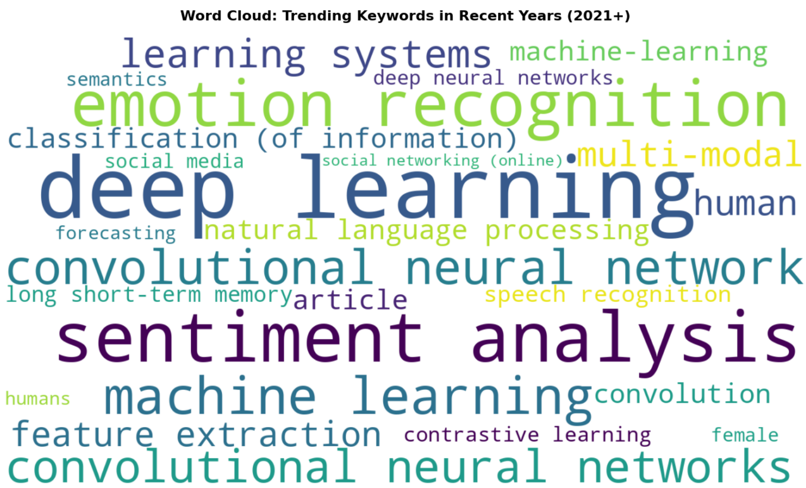

Fig. 6. Word Cloud Trending Keywords

Based on Figure 7, a total of 12 major themes were identified in the sentiment analysis research landscape. Ensemble Methods stand out as the most prominent research area with the highest prevalence of 3,675 papers with approximately 3,700 occurrences, reflecting their widespread acceptance as an effective approach for improving reliability in sentiment analysis. Transformer architecture comes in second with around 3,200 occurrences, showing how researchers have widely adopted attention-based models since 2018. What's particularly interesting is that Transformers have become the fastest-growing area, which reflects a clear shift in how the field approaches sentiment analysis. Multi-Modal approaches and Social Media Analytics have also gained significant traction (roughly 2,800 and 2,700 occurrences respectively), indicating that researchers now prefer working with diverse data sources to address real-world challenges. There's a noticeable gap between the top themes (Ensemble and Transformer) and mid-level themes (Attention Mechanism and Transfer Learning), which shows that research tends to concentrate on well-established methods. Interestingly, themes like Domain Adaptation and Low-resource Languages, while practically important, appear less frequently, indicating they remain underexplored and could offer valuable opportunities for future research.

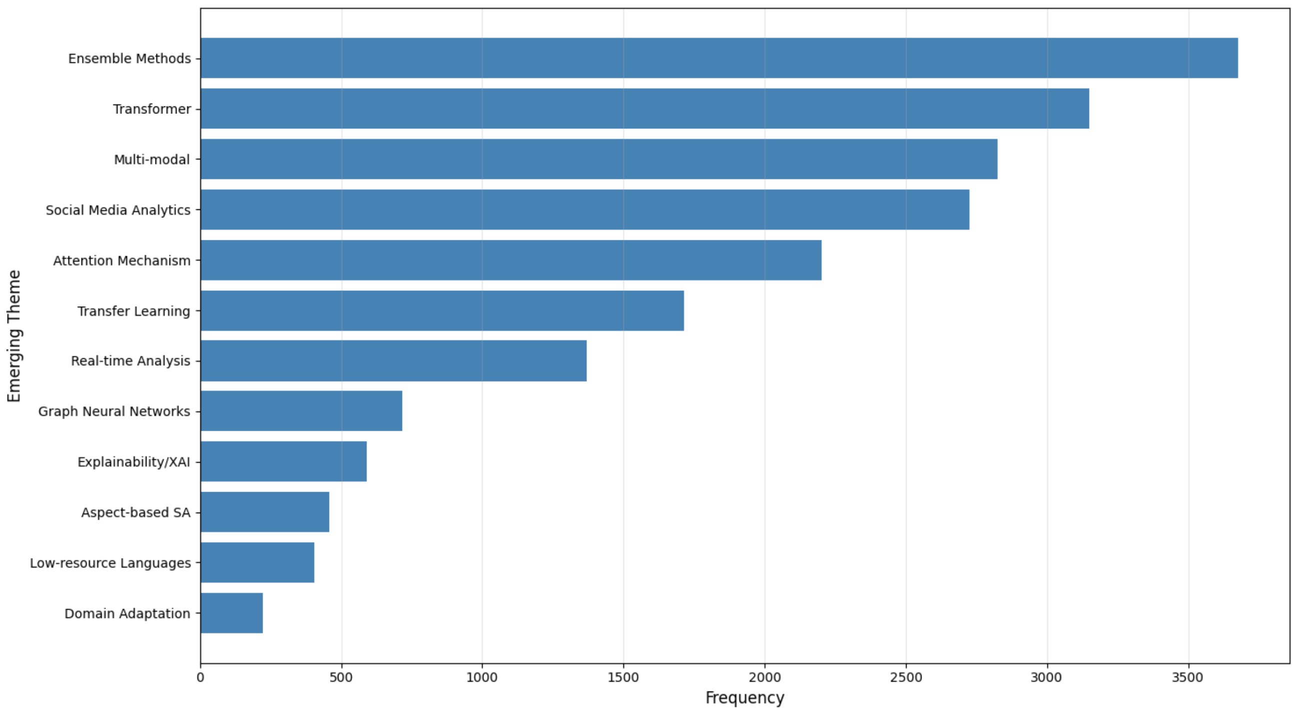

Fig. 7. Top 12 Emerging Themes in Hybrid Deep Learning in Sentiment Analysis

The co-occurrence analysis in Figure 8 reveals strategic combination patterns, with Transformer + Multi-modal with 841 papers emerging as the most prevalent pairing, indicating researchers' focus on leveraging transformer architectures for complex, multi-source sentiment analysis. Transformer's dominance is further evidenced by its appearance in 4 of the top 7 pairs, particularly with Social Media Analytics (839) and Ensemble Methods (713), suggesting its versatility across application domains. The strong co-occurrence of Attention Mechanism with both Multi-modal (701) and Transformer (643) confirms attention's role as a foundational component in modern architectures. Social Media Analytics' frequent pairing with multiple themes (appearing 4 times in top 10) reflects its position as a primary real-world testbed for novel techniques. The clustering of these combinations suggests a convergence toward hybrid approaches that integrate transformers, attention mechanisms, and ensemble strategies for robust sentiment analysis systems.

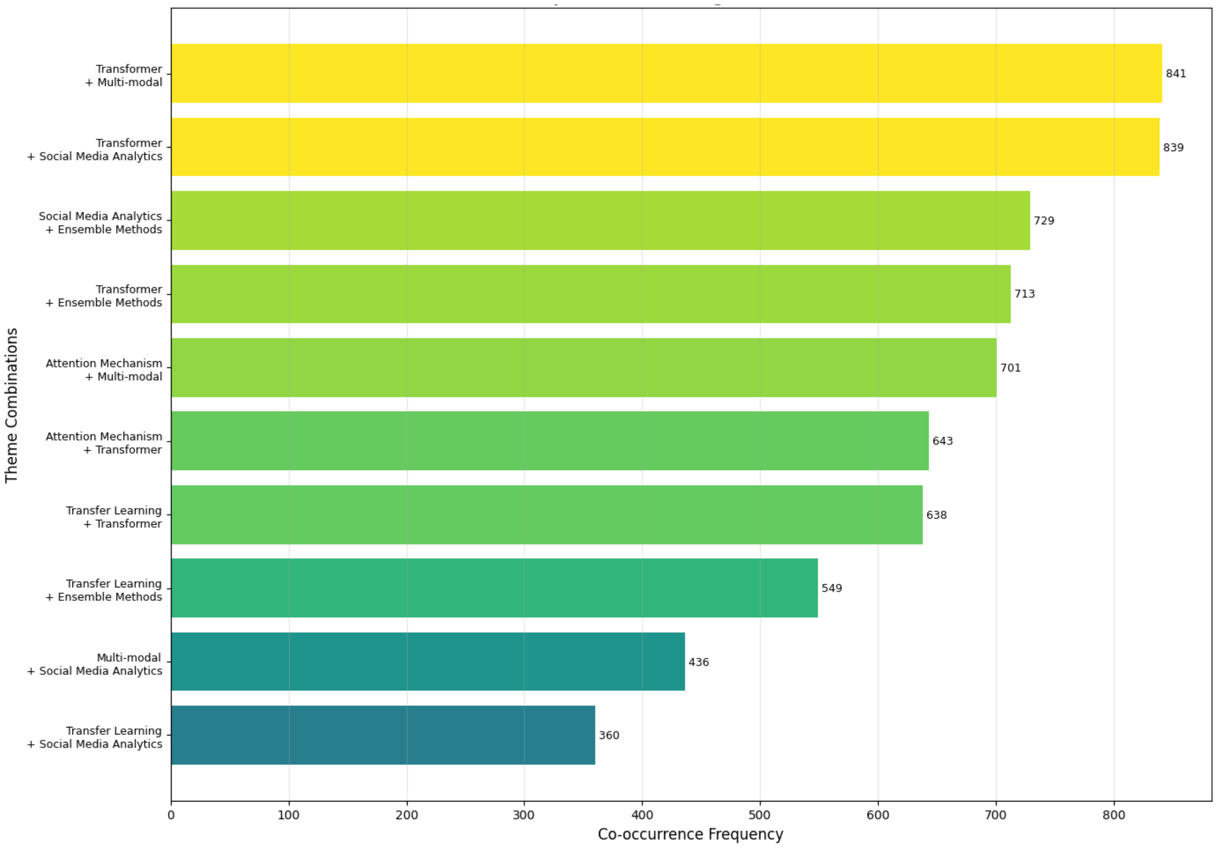

Fig. 8. Top 10 Co-occurring Themes Pairs

4.5 Novel Techniques and Architectural Innovations

Looking at Figure 9 the technique adoption timeline, the story of sentiment analysis over the past decade becomes remarkably clear. Transformer adoption surged from 2% in 2015 to 23% in 2024, reflecting how researchers tackled increasingly complex sentiment analysis challenges in today's nuanced online communication. Transformer methods grew 18.4% and Attention Mechanisms rose 12.1% because they solve what earlier models couldn't capturing long-range dependencies and contextual subtleties that help detect sarcasm, implicit sentiment, and culturally-specific expressions. BERT variants followed a different pattern; they held steady until 2021 before climbing rapidly to around 11% by 2024. Multi-head Attention hit its highest point in 2022 at 11.5% before levelling off at 10.4%, showing it has found its niche without taking over. The CNN-LSTM combination has inched up from under 1% to 3%, which confirms these hybrid approaches still have value for specific tasks. Transfer Learning has held steady between 0.5% and 1.5% surprisingly low considering how important it is, but this probably just reflects that it's become such standard practice that researchers don't even mention it anymore. The main takeaway is that Transformer-based methods have become the go-to approach, while other techniques still get chosen when they're better suited for particular problems.

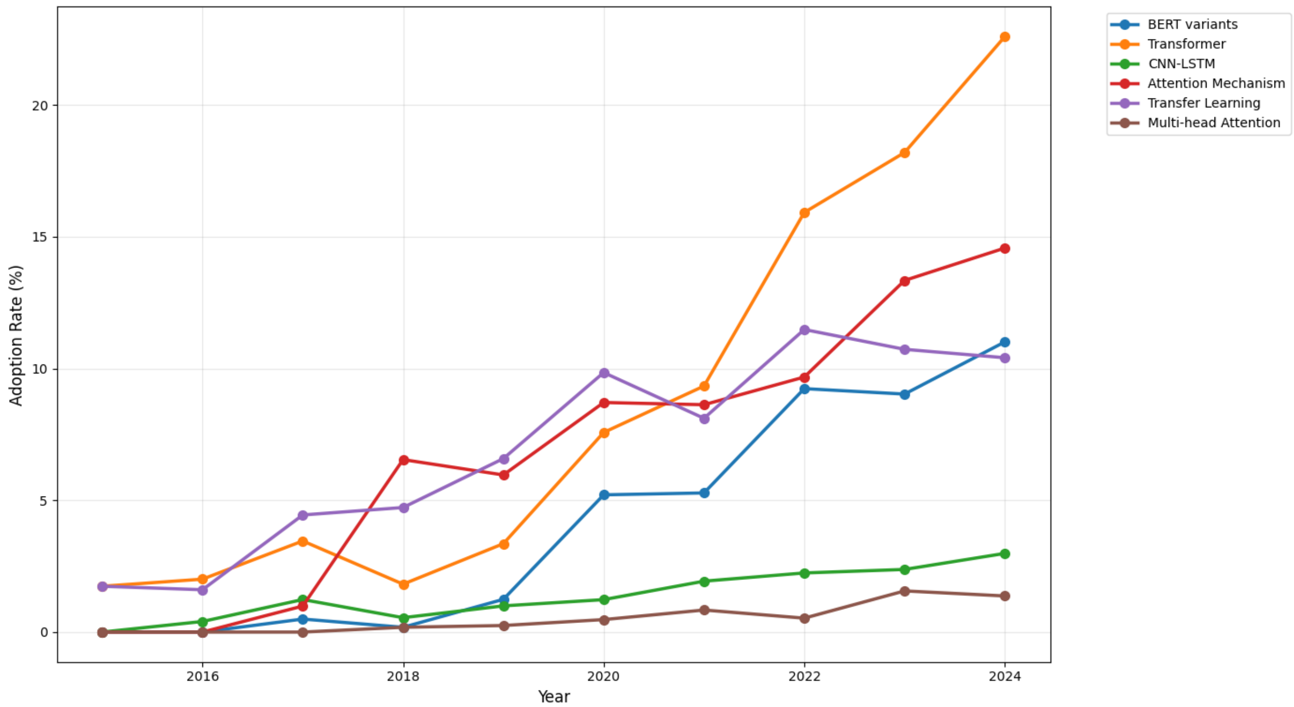

Fig. 9. Technique Adoption Timeline (2015-2024)

The comparative analysis shown in Table 6 represented data between early period (before 2018) with recent years (after 2021), sentiment analysis research has clearly changed direction. Transformer-based methods have seen the most dramatic growth, with rising from 2.4% to 20.7% (+18.4%), Attention Mechanisms growing from 1.9% to 14.0% (+12.1%), and BERT variants shooting up from almost nothing (0.1%) to 10.4% (+10.2%). This shows researchers have genuinely shifted toward attention-based models. Multi-modal approaches and Social Media Analytics have also grown significantly, up 9.3% and 8.1% respectively, which means more people are working with mixed data types and messy real-world content instead of clean test datasets. Transfer Learning's 7.7% growth (from 2.7% to 10.4%) demonstrates how pre-trained models have become standard practice in the field. What's curious is that Ensemble Methods, even though they're still the most common overall, only grew 3.9% (from 16.6% to 20.5%)—they seem to have reached a stable point rather than continuing to expand. Techniques in the middle like Graph Neural Networks (+4.0%), Explainability/XAI (+3.5%), and Real-time Analysis (+2.4%) are picking up steam gradually, while newer areas like Low-resource Languages, Aspect-based SA, and CNN-LSTM each grew about 2%. The smallest changes showed up in specific BERT hybrids like BERT-CNN (+0.2%) and BERT-LSTM (+0.1%), and T5/BART stayed flat, which suggests mixing BERT with older methods hasn't really paid off.

Table 6
Comparative Analysis: Early vs Recent Period

Pattern	Early (≤2018)	Recent (≥2021)	Change	Trend
Transformer	2.4%	20.7%	+18.4%	📈
Attention Mechanism	1.9%	14.0%	+12.1%	📈
BERT variants	0.1%	10.4%	+10.2%	📈
Multi-modal	7.8%	17.1%	+9.3%	📈
Social Media Analytics	7.8%	15.9%	+8.1%	📈
Transfer Learning	2.7%	10.4%	+7.7%	📈

Graph Neural Networks	0.7%	4.7%	+4.0%	→
Ensemble Methods	16.6%	20.5%	+3.9%	→
Explainability/XAI	0.4%	3.9%	+3.5%	→
Real-time Analysis	5.6%	8.0%	+2.4%	→
Low-resource Languages	0.4%	2.5%	+2.1%	→
Aspect-based SA	0.7%	2.8%	+2.1%	→
CNN-LSTM	0.4%	2.6%	+2.1%	→
CNN-BiLSTM	0.0%	0.9%	+0.9%	→
GPT models	0.0%	0.8%	+0.8%	→
XLNet	0.0%	0.4%	+0.4%	→
Domain Adaptation	1.0%	1.3%	+0.3%	→
BERT-CNN	0.0%	0.2%	+0.2%	→
BERT-LSTM	0.0%	0.1%	+0.1%	→
T5/BART	0.2%	0.2%	-0.0%	→

These findings address a critical need in sentiment analysis research by identifying which hybrid architectures and methods researchers are actually adopting. The popularity of adaptable ensemble methods, the rapid rise of transformer architectures, and the growing focus on multi-modal and social media analytics indicate that the field is moving away from isolated model development toward integrated systems that handle real-world complexity.

Despite these advances, major gaps remain in the research. Low-resource languages, domain adaptation, and real-time processing get far less attention than they deserve. Targeting these areas could dramatically improve sentiment analysis accessibility, accuracy, and global reach. These patterns give researchers a roadmap to align their work with both cutting-edge methodological advances and overlooked practical needs.

5. Limitations

This bibliometric analysis outlines several limitations that should be considered when interpreting the findings.

i. Data Source Limitations: Exclusive reliance on the Scopus database means that the study may not include research published in non-indexed conferences, preprint servers, or regional journals, particularly works in languages other than English.
ii. Search Strategy Limitations: Although the search query undergoes careful design, it may have overlooked relevant studies that employed different terminology or did not explicitly describe their methods using terms like "hybrid models," "ensemble," "combine," or "fusion."
iii. Methodological Limitations: Citation-based metrics tend to privilege older publications that have gathered more citations through time, which may undervalue recent innovative work. The analysis lacks the ability to separate positive citations that extend previous research from critical citations that question earlier findings. Furthermore, keyword analysis relies on terms chosen by the authors themselves, which may not fully represent the actual methodological approach or contributions contained in their studies.
iv. Temporal Constraints: The fifteen-year timeframe may understate very recent innovations still accumulating citations. Despite these constraints, the study still offers valuable insights into how research has developed, which contributions have shaped the

field, and how themes have evolved in hybrid deep learning approaches for sentiment analysis.

6. Conclusion

This bibliometric study investigated how sentiment analysis research with hybrid deep learning models has progressed from 2010 to 2025, analysing 18,373 papers across five main research objectives.

Responding to RQ1 regarding publication and citation trends, scholars have demonstrated striking growth in sentiment analysis research during the study period. Authors generated only 86 papers in 2010, yet by 2025, the community had expanded its productivity to 4,008 papers, multiplying output almost 47 times. Researchers in this field have accumulated 266,850 citations for their work, generating an average of 17,790 citations each year. These figures highlight how significantly the field influences academic discourse and how consistently scholars engage with this research, showing that sentiment analysis continues to flourish as researchers pursue innovative approaches and applications in hybrid deep learning methodologies.

Addressing RQ2 on dominant hybrid approaches, researchers have favoured Hybrid-General or Ensemble Methods most heavily, applying them across 1,842 studies. Out of the studies examined, 1,087 (21%) employed CNN-LSTM combinations and 818 (16%) used Attention-LSTM techniques. These three approaches dominate the field because scholars find them reliable for sentiment analysis work. The pattern here is clear: researchers prefer flexible, modular setups that let them swap components in and out based on their specific data and research questions.

For RQ3 on major application domains, researchers identified 12 major themes, with Emotion Recognition leading at 21.9% by 2024 which it up from just 11.9% in 2015-2017. Multi-modal approaches (20.6%) and Social Media Analytics (15.9%) also saw strong increases of 9.3% and 8.1% respectively. These numbers confirm that researchers have shifted their focus toward real-world challenges involving diverse data sources instead of sticking with controlled benchmark datasets.

On RQ4 concerning temporal evolution of emerging themes, the field went through distinct phases: traditional methods dominated early on (2015-2017), transformers-built momentum in the middle years (2018-2020), and sophisticated neural architectures became standard practice recently (2021-2024). Everything accelerated dramatically after 2021, coinciding with GPT-3 and ChatGPT's release, which showcased what transformers could achieve and prompted researchers everywhere to adopt them.

Regarding RQ5 about novel techniques and architectural innovations, researchers adopted Transformer variants went from 2.4% to 20.7% (+18.4%), accelerating mainly between 2020-2024. Attention Mechanisms grew from 1.9% to 14.0% (+12.1%), while BERT variants jumped from nearly invisible (0.1%) to 10.4% (+10.2%). These figures demonstrate that attention-based architectures became a leading approach in the field, though methodological diversity persists with ensemble methods (20.5%) and traditional hybrids remaining popular for specific applications.

The study also revealed important gaps: Low-resource Languages, Domain Adaptation, and Aspect-based Sentiment Analysis each account for less than 3% of recent research, despite their real-world importance. These overlooked areas offer promising opportunities for future work, particularly in making sentiment analysis accessible across different languages and specialized fields.

Publication trends show researchers moving from traditional machine learning to transformer-based architectures, with adoption rates increasing from 2.4% to 20.7%. Researchers have also expanded their efforts toward multi-modal approaches (17.1%), social media analytics (15.9%), and explainability/XAI (3.9%). However, the data exposes underexplored areas, with low-resource languages growing only 2.1%, domain adaptation just 0.3%, and real-time analysis 2.4%. This publication gap suggests researchers have concentrated on advancing model sophistication while giving less attention to accessibility challenges. The limited research results in this practical area indicate potential opportunities for future work, particularly in developing solutions that extend sentiment analysis capabilities beyond high-resource languages and well-studied domains.

Acknowledgement
This research was not funded by any grant.

References

[1] Zhang, Lei, Shuai Wang, and Bing Liu. "Deep learning for sentiment analysis: A survey." *Wiley interdisciplinary reviews: data mining and knowledge discovery* 8, no. 4 (2018): e1253. https://doi.org/10.1002/widm.1253

[2] Yadav, Ashima, and Dinesh Kumar Vishwakarma. "Sentiment analysis using deep learning architectures: a review." *Artificial Intelligence Review* 53, no. 6 (2020): 4335-4385. https://doi.org/10.1007/s10462-019-09794-5

[3] Liu, Bing, and Lei Zhang. "A Survey of Opinion Mining and Sentiment Analysis." Mining Text Data, 2012, 415–63. https://doi.org/10.1007/978-1-4614-3223-4_13

[4] Pang, Bo, and Lillian Lee. "Opinion mining and sentiment analysis." *Foundations and Trends® in information retrieval* 2, no. 1–2 (2008): 1-135. https://doi.org/10.1561/1500000011

[5] Vinodhini, G., and R. M. Chandrasekaran. "Sentiment analysis and opinion mining: a survey." *International Journal* 2, no. 6 (2012): 282-292.

[6] Mohammad, Saif M. "Challenges in sentiment analysis." In *A practical guide to sentiment analysis*, pp. 61-83. Cham: Springer International Publishing, 2017. https://doi.org/10.1007/978-3-319-55394-8_4

[7] Chen, Yuanxing, Qingzhao Zhang, Shuangge Ma, and Kuangnan Fang. "Heterogeneity-aware clustered distributed learning for multi-source data analysis." *Journal of Machine Learning Research* 25, no. 211 (2024): 1-60.

[8] Sofi, Shakirah Mohd, and Ali Selamat. "Aspect based sentiment analysis: feature extraction using latent dirichlet allocation (LDA) and term frequency-inverse document frequency (TF-IDF) in machine learning (ML)." *Malaysian Journal of Information and Communication Technology (MyJICT)* (2023): 158-168. https://doi.org/10.53840/myjict8-2-102

[9] Zhang, Ye, and Byron C. Wallace. "A sensitivity analysis of (and practitioners' guide to) convolutional neural networks for sentence classification." In *Proceedings of the Eighth International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pp. 253-263. 2017.

[10] S. Minaee, N. Kalchbrenner, E. Cambria, N. Nikzad, M. Chenaghlu, and J. Gao, "Deep Learning-Based Text Classification," Apr. 30, 2022, *Association for Computing Machinery*. https://doi.org/10.1145/3439726

[11] Sherstinsky, Alex. "Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network." *Physica D: Nonlinear Phenomena* 404 (2020): 132306. https://doi.org/10.1016/j.physd.2019.132306

[12] Tay, Yi, Mostafa Dehghani, Dara Bahri, and Donald Metzler. "Efficient transformers: A survey. arXiv." *arXiv preprint arXiv:2009.06732* (2020). https://doi.org/10.1145/3530811

[13] M. E. Basiri, S. Nemati, M. Abdar, E. Cambria, and U. R. Acharya, "ABCDM: An Attention-based Bidirectional CNN-RNN Deep Model for sentiment analysis," *Future Generation Computer Systems*, vol. 115, pp. 279–294, Feb. 2021, doi: 10.1016/j.future.2020.08.005.

[14] Devlin, Jacob, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. "Bert: Pre-training of deep bidirectional transformers for language understanding." In *Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technologies, volume 1 (long and short papers)*, pp. 4171-4186. 2019. https://doi.org/10.18653/v1/n19-1423

[15] Liu, Yinhan, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. "Roberta: A robustly optimized bert pretraining approach." *arXiv preprint arXiv:1907.11692* (2019).

[16] Sun, Chi, Luyao Huang, and Xipeng Qiu. "Utilizing BERT for aspect-based sentiment analysis via constructing auxiliary sentence." *arXiv preprint arXiv:1903.09588* (2019).

[17] Xu, Hu, Lei Shu, Philip Yu, and Bing Liu. "Understanding Pre-Trained BERT for Aspect-Based Sentiment Analysis." Proceedings of the 28th International Conference on Computational Linguistics, 2020, 244–50. https://doi.org/10.18653/v1/2020.coling-main.21

[18] Gao, Ziwen, Zhiyi Li, Jiaying Luo, and Xiaolin Li. "Short text aspect-based sentiment analysis based on CNN+ BiGRU." *Applied Sciences* 12, no. 5 (2022): 2707. https://doi.org/10.3390/app12052707

[19] Keya, Ashfia Jannat, Hasibul Hossain Shajeeb, Md Saifur Rahman, and M. F. Mridha. "FakeStack: Hierarchical Tri-BERT-CNN-LSTM stacked model for effective fake news detection." *Plos one* 18, no. 12 (2023): e0294701. https://doi.org/10.1371/journal.pone.0294701

[20] Adoma, Acheampong Francisca, Nunoo-Mensah Henry, and Wenyu Chen. "Comparative analyses of bert, roberta, distilbert, and xlnet for text-based emotion recognition." In *2020 17th international computer conference on wavelet active media technology and information processing (ICCWAMTIP)*, pp. 117-121. IEEE, 2020. https://doi.org/10.1109/iccwamtip51612.2020.9317379